IN THE MATTER of European Patent No. 0 125 023 of Genentech, Inc.

-and-

IN THE MATTER of an Opposition thereto filed by Celltech Limited

STATUTORY DECLARATION

I, Geoffrey Thomas YARRANTON of 8, Sherwood Road, Winnersh, Near Wokingham, Berkshire, U.K., do solemnly and sincerely declare as follows:

1. I am at present employed by Celltech Limited as Deputy Director of Research. I have held this position for the past two years. I attach as Exhibit GTY1 hereto my *curriculum vitae* which shows my academic qualifications and my employment record. I attach as Exhibit GTY2 a list of publications which I have authored or co-authored and patent applications in which I am named as the inventor or a co-inventor. I attach as Exhibit GTY3 a copy of each of the publications to which I refer in this Declaration.

2. I have been working in the field of genetic engineering since 1979 and have been concerned in particular with recombinant gene expression systems.

3. I have read UK Patent Application No. 8308235 of Boss *et al* (the Boss application) and understand its content. I was employed by Celltech at the time of the filing of the first application covering the subject matter in March 1983. I was aware of the background to the invention, although I was not involved in the making of the invention.

4. In March 1983, I was well aware of the body of knowledge developed in relation to antibodies and the gene system involved in their production in mammals. This body of knowledge, particularly in relation to the gene system, had been considerably expanded following the development of the Kohler-Milstein technique for the production of monoclonal antibodies in 1975 and by use of the techniques of recombinant DNA technology thereafter.

5. By March 1983, it was well understood that an antibody molecule comprises two identical heavy chains and two identical light chains arranged as shown in Figure 1 of the Boss application. By various techniques, in particular by X-ray crystallography, it had been shown that each light chain comprised two distinct regions of amino acids which folded into a recognisable structure. These distinct regions, which are generally called domains, are connected together by a short, relatively unfolded sequence of amino acids. Similarly, the heavy chain consists of four or five domains (depending on the class of antibody) linked by relatively unfolded sequences. It was also known that between the second and third domains in the heavy chain there is a flexible area known as the hinge region.

6. It was also well known that the N-terminal domains of each heavy/light chain pair fold together to form an antigen binding site. Investigations using X-ray crystallography and protein sequence analysis showed that in the N-terminal domains of both the heavy and the light chain there are three areas which appear to be responsible for binding the antibody to its antigen. These areas have a very variable amino acid sequence and for this reason the N-terminal domains are generally referred to as variable domains. The areas with very variable sequence are generally referred to as hypervariable regions or complementarity determining regions (CDRs). The remaining domain or domains in the heavy and light chains are known as constant domains, since their amino acid sequences do not vary significantly for each class of antibody.

7. It had also been known for some time before the date of filing of the Boss application that it is possible to produce antibody fragments by enzymic digestion of whole antibody molecules. For instance, it is possible to digest antibody molecules in such a way as to cleave the two heavy chains on the C-terminal side of the hinge region, leading to the production of a bivalent antigen-binding fragment known as the $F(ab')_2$ fragment. Alternatively, the heavy chains can be cleaved on the N-terminal side of the hinge region to produce two monovalent antigen-binding fragments, known as the Fab fragments, and a fragment known as the Fc fragment. Moreover, by careful digestion of Fab fragments, it is possible to cleave between the first and second domains in each chain to produce monovalent antigen-binding fragments known as Fv fragments.

8. It had been well known for many years before the date of filing of the Boss application that a single mammal can produce antibodies against a large number of antigens. Before the advent of genetic engineering, there were two theories as to how it was possible for this to be achieved. These were known as the "germ line" and the "somatic" theories of antibody diversity.

9. According to the germ line theory, the DNA in each germ (sperm or egg) cell contained a very large number of genes, each encoding a different variable domain. When challenged by an antigen, the mammal's immune system selected the appropriate gene from the repertoire and used this gene to synthesise an antibody which would recognise the antigen.

10. According to the somatic theory, the germ cells contained a relatively small number of genes, each encoding a different prototype variable domain. On challenge by an antigen, the appropriate prototype gene was selected and then mutated so that the antibody synthesised from the mutated prototype gene had the required specificity.

11. With the advent of recombinant DNA technology, it became possible to determine which of these theories was most nearly correct and to determine how that theory needed to be amended to explain

antibody diversity. By March 1983, these investigations into the antibody gene system had progressed a long way and it was possible to explain in principle how antibody diversity was generated.

12. As regards the light chain, it was shown that in the germ line, there are several hundred genes which encode most, but not all, of the variable domain. These genes encode from the N-terminal end nearly to the C-terminal end of the variable domain. These genes encode part, but not all, of the third CDR. These genes are called V genes and generally comprise an intron separating a leader sequence, encoding a signal peptide, from the variable region coding sequence. There are also about five genes which encode the remainder of the C-terminal end of the variable domain and the amino acid sequence which joins the variable domain to the constant domain. These genes are called the J (for "joining") genes. (There are also genes, the C genes, which encode each constant domain, but variability in the constant domains does not affect antibody diversity. These C genes are therefore not discussed in detail at present).

13. During B cell development, there is a rearrangement of the genes within the germ line to produce an active gene which encodes the desired light chain. It was believed in March 1983 that, in the rearrangement, one of the V genes is recombined with one of the J genes and the appropriate C gene to form a rearranged gene comprising, from the 5' end, a leader sequence, an intron, the V gene, the J gene, an intron and the C gene. After translation to produce RNA, the introns are removed by splicing to produce mRNA comprising the leader, V, J and C sequences fused together.

14. In March 1983, it was believed that the situation for the heavy chain is similar, but a little more complicated. In addition to the V and J region genes, the heavy chain gene system in the germ line also includes a number of genes called D genes. In this case D stands for "diversity". Again, the V genes encode from the N-terminal end almost to the C-terminal end of the variable domain, but do not encode all of the third CDR. The D genes encode a further part of the CDR. The remainder of the third CDR and the sequence

joining the variable domain to the first constant domain is encoded by the J gene.

15. In heavy chains, there are more types of constant region. In the rearranged form, each constant region is encoded by a different gene system. A typical rearranged gene system, as it was understood in 1983, is shown in Figure 1 attached as Exhibit GTY4 hereto. This shows that each gene system comprises exons encoding each domain and the hinge region, each exon being separated from the adjacent exons by an intron.

16. Rearrangement to produce a heavy chain gene during B cell development involves a selected V gene being fused to a selected D gene which in turn is fused to a selected J gene. There then follows an intron which leads into the gene system encoding the constant domains. During mRNA production, the introns are removed in the normal manner.

17. It can thus be seen that antibody diversity is generated by the way in which the selected V, J and, in the case of heavy chains, D genes are recombined from the germ line to produce the active genes. It was also thought at that time that further diversity was generated by a small degree of mutation occurring during rearrangement.

18. The body of knowledge on the generation of antibody diversity had before 1983 been gathered together in a number of review articles. A relatively simple review article was presented in Scientific American, 246, 5, 73-83, 1982 by P. Lederer. A more detailed account is given in Genetic Engineering, 3, 157-188, 1981 by P. Early and L. Hood. Among the papers which contributed to this body of knowledge are: PNAS-USA, 74, 8, 3518-3522, 1977 (Tonegawa *et al*); Nature, 280, 370-375, 1979 (Seidman *et al*); Nuc. Acids Res., 8, 8, 1709-1720, 1980 (Steinmetz *et al*); Nuc. Acids Res., 8, 3591-3601, 1980 (Zakut *et al*); and Cell. 19, 981-992, 1980 (Early *et al*).

19. The structure of the constant region gene system had been disclosed in such review articles as Ann. Rev. Immunol., 1, 393-422, 1983 (Wall and Kuehl). Among the papers which contributed to the body of knowledge in this respect are: Cell, 22, 197-207, 1980 (Hieter *et al*); DNA, 1, 1, 11-18, 1981 (Ellison *et al.*); and Nuc. Acids Res., 10, 13, 4071-4079, 1982 (Ellison *et al.*)

20. It can thus be seen that by March 1983 a large amount was known about the arrangement of the antibody gene system in the germ line, what the active gene looked like after rearrangement and what the mRNA encoding the heavy and light chains looked like before translation. Moreover, the DNA sequences for constant regions in at least the human and mouse genomes were known, for instance from the Hieter and two Ellison references referred to in the preceding paragraph. In particular, it was in 1983 believed that in an active, rearranged light chain gene there is a leader sequence, a first intron, a sequence encoding the complete variable domain and joining sequence, a second intron and a sequence encoding the complete constant domain. Similarly, it was in 1983 believed that an active, rearranged heavy chain gene has the same structure as a light chain gene except that, following the sequence encoding the first constant domain, there is a third intron, a sequence encoding the hinge region, a fourth intron, a sequence encoding the second constant domain, a fifth intron, a sequence encoding the third constant domain and, in the case of some classes of antibody, a sixth intron and a sequence encoding the fourth constant domain. Such a rearranged heavy chain gene is shown in Figure 1.

21. By March 1983, it was known that the immunoglobulin gene systems for all mammalian species studied are organised in a similar manner. In particular, the mouse and human immunoglobulin gene systems have similar organisations both in the germ line and as rearranged. This was shown, for instance, by comparisons of mouse and human genomic DNA sequences encoding constant regions. In this respect, I refer to the 1982 Ellison *et al*. paper referred to above.

22. Although a large amount was known about the antibody gene system, in March 1983 much remained unknown. In particular, the mechanisms whereby the appropriate V, D, J and C genes are selected in response to antigen stimulation, the mechanisms by which the genes are rearranged, and the way in which the light and heavy chains, once synthesised, are assembled into complete antibody molecules and exported from the cells had not been elucidated.

23. Turning to the Boss application, it is stated on page 7, lines 7 to 12 that the inventors contemplated that "chimeric" antibodies could be produced following the teaching in the application. By "chimeric" antibodies, I mean an antibody wherein the variable domains in the heavy and light chains are derived from an antibody of a first class or origin and the constant domains are derived from an antibody of a different class or origin. I believe that, at that time, any competent recombinant DNA technologist, given the general body of knowledge concerning the structure of the antibody gene system and the complete DNA sequence information, would have been able to construct a gene encoding a chimeric antibody or a chimeric antibody fragment without undue effort.

24. In the Boss application, it is stated on page 7, lines 7 to 12 that antibody molecules may be produced by combining the variable domains from one antibody with the constant region from another antibody. Any competent recombinant DNA technologist would have been able to do this without any undue difficulty. All that he would have needed was the knowledge of the structure and DNA sequences of the antibody gene system and knowledge of the conventional techniques in the recombinant DNA art. He would then have been able readily to determine the sequences encoding the variable domain from one source and the constant domain (for a light chain) or the constant region (for a heavy chain) from another source and fuse these together to form a gene encoding a chimeric heavy or light chain. Techniques for carrying out these manipulations were well known, for instance from Maniatis (Molecular Cloning, A Laboratory Manual, Maniatis et al., Cold Spring Harbour Laboratory, 1982) which was at that time the "bible" for recombinant DNA technologists.

25. For instance, a chimeric heavy chain gene could have been produced as shown in either of the two Examples given in Exhibit GTY5 attached hereto. The first Example shows the production of a chimeric heavy chain gene by site directed mutagenesis. The second Example shows the use of total chemical synthesis. Both the Examples use only information derived from published papers and techniques well known in the art in early 1983. It would also have been feasible to produce an Example showing the use of mRNA extraction, cDNA production and manipulation of the cDNA to produce a chimeric heavy chain gene. I chose to illustrate the routes using total chemical synthesis and site directed mutagenesis since these allow the genes to be constructed directly from published DNA sequence information.

26. By March 1983, it was also known that antibody fragments comprise various combinations of domains. An Fv fragment comprises the variable domains of the heavy and light chain. A Fab fragment comprises the complete light chain and the variable and first constant domains of the heavy chain. A $F(ab')_2$ fragment comprises the same domains as a Fab fragment and additionally includes the hinge region on the C-terminal end of the first constant domain in the heavy chain. The presence of the hinge region allows the fragment to be present as a divalent product since the two antigen-binding portions are held together by cross-linking through disulphide groups in the hinge region.

27. From the body of knowledge on the structure of rearranged antibody genes and in particular from the literature attached as Exhibit GTY3 it is readily possible to determine where, in the rearranged gene for any particular antibody, the sequence encoding each domain or the hinge region ends and where the next intron begins. It is then a matter of routine manipulation to construct a cDNA gene with deletion of undesired sequences from a genomic gene. Thus, genes encoding chimeric antibody fragments could have been produced by a person of ordinary skill in the art without undue effort, for instance by following a procedure similar to that described in paragraph Exhibit GTY5.

28. As shown above, the manipulations required to produce the appropriate active gene encoding a chimeric antibody or fragment are routine for a competent recombinant DNA technologist, given the body of knowledge about the structure of rearranged antibody genes. If this is taken in conjunction with the teaching in the Boss application, I would have expected such a person to be able to follow the teaching in the Boss application and produce chimeric antibodies or fragments by expression of appropriately constructed genes in a separate host cells or by coexpression of both chains in a single host cell. If the skilled person had not had the teaching in the Boss application, I believe that he would not have even contemplated trying to produce any sort of antibody molecule by coexpression in a single host cell.

AND I MAKE THIS SOLEMN DECLARATION sincerely believing the same to be true and by virtue of the Statutory Declarations Act 1835.

Signed by.....
Geoffrey Thomas YARRANTON

This..... day of..... 1992

10

IN THE MATTER of
European Patent No. 0 125 023
of Genentech, Inc.
-and-
IN THE MATTER of an
Opposition thereto filed by
Celltech Limited

I hereby certify that this is Exhibit GTY1 to the Statutory Declaration of Geoffrey Thomas YARRANTON executed this day of 1992.

CURRICULUM VITAE

NAME: Geoffrey Thomas Yarranton

DATE OF BIRTH: 12.11.51.

MARITAL STATUS: Married

NATIONALITY: British

PLACE OF BIRTH: Burton-on-Trent, Staffs. England

ADDRESS: Home: 8 Sherwood Road,
Winnersh,
Nr. Wokingham, Berks.
Telephone: 0734-784129

Work: Dept. of Molecular Genetics,
Celltech Ltd.,
216 Bath Road,
Slough, Berks. SL1 4EN.
Telephone: 0753 34655

EDUCATION:

Date	School/Institute	Qualifications	Class
1963-1970	Burton-on-Trent Grammar School	10 'O' levels 3 'A' levels	
1970-1973	University of Leicester School of Biological Science	B.Sc.	1st Class (Hons)
1973-1976	National Institute for Medical Research, Genetics Division.	Ph.D.	
1989	Ashridge Management College	G-course.	

EMPLOYMENT:

Date	Institute	Position
1976-1979	Massachusetts Institute of Technology Department of Biology	Jane Coffin Childs Post-Doctoral Fellow
1979-1982	National Institute for Medical Research, Genetics Division	Research Scientist
1982-1984	Celltech Ltd.	Group Leader - Microbial Genetics
1984-1986	Celltech Ltd.	Head of Department - Microbial Genetics
1986-1989	Celltech Ltd.	Head of Department - Molecular Genetics
1989-1990	Celltech Ltd.	Sub Divisional Head - Molecular Genetics

Studies in the prokaryote, E. coli, involved the 'SOS' response which is caused by the co-ordinate induction of the genes forming the 'SOS' regulon. Induction is triggered by agents which damage DNA. Expression of the inducible genes is controlled at the transcriptional level by a repressor protein lexA, and by an 'inducer' protein recA. The recA protein is a multifunctional protein, acting both to induce the genes of the regulon, and to act directly in DNA repair/recombination. Using In vitro recombinant DNA technology and In vivo transposition mutagenesis, the functions of recA in regulatory and direct catalytic processes were dissected out. The approach used, made use of negative complementation, a technique which requires the expression of protein subunits which interfere with the activity of 'normal' subunits.

Evidence was also obtained for the presence of inducible DNA repair rocesses in the simple eukaryotes Saccharomyces cerevisiae and Ustilago maydis.

Present Job Responsibilities

Research at Celltech

My research at Celltech has been concerned with (1) the development of stable, heterologous gene expression systems, (2) the engineering of monoclonal antibodies for human therapy and (3) the identification and selection of new therapeutic targets. An outline of my work in the first two categories is given below.

1. Bacterial Expression Vectors : Dual-Origin Vectors

   Many expression vectors designed for heterologous gene expression in E. coli have been found to be segregationally unstable, particularly when used in large scale fermentations. One of the major determinants in this instability is the translation of heterologous mRNA rather than transcription, even though the heterologous gene product itself may not be toxic. To overcome the problem of plasmid instability whilst retaining high level gene expression, a novel copy number-controlled vector was developed. The dual-origin comprises two origins of replication, one expressed constitutively defining a low copy number (5 copies per cell) and the other expressed conditionally, giving a copy number of 200-300 copies per cell.

   The dual origin vectors have been studied in chemostat culture as well as batch culture and are stable, in the absence of antibiotic selection, for over 150 generations. A variety of heterologous proteins have now been expressed in E. coli using this vector system and scale-up above 150 litres has been demonstrated.

   Recently, the ability of E. coli to secrete some heterologous proteins either into the periplasm or medium has been exploited. Expression vectors employing E. coli signal sequences or the C-terminal haemolysin signal have been evaluated for the production of soluble, active heterologous proteins.

2. S. cerevisiae Gene Expression and Protein Secretion

   S. cerevisiae is a simple eukaryote, widely used in the food industry and in academic research laboratories. Because of problems associated with the insolubility of heterologous proteins made in E. coli, considerable effort has been made to use S. cerevisiae, both for intracellular protein production and protein secretion. In

Administrative responsibilities

The Molecular Genetics Sub-Division comprises three departments:- Protein Engineering, Gene Cloning and Mammalian Gene Expression. There are forty research scientists in the Sub-Division approximately 70% of whom possess a PhD degree. Work within the Sub-Division includes contract research and own-risk discovery projects.

Line management duties consist of recruitment, staff appraisal and career planning, scientist management, safety monitoring and budgetary control.

In addition to Sub-Divisional responsibilities I chair two groups, these are concerned with (1) Maintaining scientific standards on ongoing projects (Science Review Group) and (2) initiating and progressing new ideas at the pre-project stage (Innovation Support Group). The Science Review Group consists of all Celltech senior scientists and is the major review group within Research and Development.

11

IN THE MATTER of
European Patent No. 0 125 023
of Genentech, Inc.
-and-
IN THE MATTER of an
Opposition thereto filed by
Celltech Limited

I hereby certify that this is Exhibit GTY2 to the Statutory Declaration of Geoffrey Thomas YARRANTON executed this day of 1992.

Spanos, A., Sedgwick, S.G., Yarranton, G.T., Hubscher, U. and Banki, G.R. (1981). Detection of the catalytic activities of DNA polymerase and their associated exonucleases following SDS-polyacrylamide gel electrophoresis. Nuc. Acids Res. 9, 1825-1839.

Lee, M.G. and Yarranton, G.T. (1982). Inducible DNA repair in Ustilago maydis. Mol. gen. Genet. 185, 245-250.

Sedgwick, S.G., Yarranton, G.T. (1982). Cloned truncated recA genes in E. coli I. Effect on radio sensitivity and recA$^+$ dependant processes. Mol. gen. Genet. 185, 93-98.

Yarranton, G.T. and Sedgwick, S.G. (1982). Cloned truncated recA genes in E. coli II. Effects of truncated gene products on in vivo recA$^+$ protein activity. Mol. gen. Genet., 185, 99-104.

Robinson, M.K., Lilley, R., Little, S., Emtage, J.S., Yarranton, G.T., Stephens, P., Millican, A., Eaton, M. and Humphreys, G. (1984). Codon usage can affect efficiency of translation of genes in Escherichia coli. Nucleic Acids Res. 12, 6663-6671.

Yarranton, G.T., Wright, E.M., Robinson, M.K. and Humphreys, G.O. (1984). Dual-origin plasmid vectors whose origin of replication is controlled by the coliphage lambda promoter P_L. Gene 28, 293-300.

Caulcott, C.A., Lilley, G., Wright, E.M., Robinson, M.K. and Yarranton, G.T. (1985). Investigation of the Instability of Plasmids Directing the Expression of Met-Prochymosin in Escherichia coli. J. Gen. Micro. 131, 3355-3365.

Wright, E.M., Humphreys, G.O. and Yarranton, G.T. (1986). Dual-origin plasmids containing an amplifiable ColE1 ori; temperature-controlled expression of cloned genes. Gene 49, 311-321.

Bodmer, M.W., Angal, S., Yarranton, G.T., Harris, T.J.R., Lyons, A., King, D.J., Pieroni, G., Riviere, C., Verger, R. and Lowe, P.A. (1987). Molecular cloning of a human gastric lipase and expression of the enzyme in yeast. Biochem. et Biophys. Acta. 909. 237-244.

King, D.J., Walton, F., Smith, B.W., Dunn, M. and Yarranton, G.T. (1987). The Recovery of Recombinant Proteins from Yeast. Biochem. Soc. Trans. (In press).

Wright, E.M. and Yarranton, G.T. (1987). Single Point Mutations at the 5 End of a Eukaryotic Gene Can Affect Synthesis of Gene Product in E. coli. In Genetics of Translation, NATO 51 Series 14 331-342.

Yarranton, G.T. Humphreys, G.O., Robinson, M.K., Caulcott, C.A., Wright, E.M. DNA vectors and their use in recombinant DNA technology (1987). U.K. Patent GB2136814B.

Field, H., Yarranton, G.T. and Rees, A.R. (1988). A Functional Recombinant Immunoglobulin Variable Domain from Polypeptides produced in Escherichia coli. In: "Vaccines 88: Eds: Ginsbery, H., Brown, F., Lerner, R.A. and Chanock, R.M., Publ. Cold Spring Harbor, New York, pp29-34.

LIST OF PATENT APPLICATIONS

EP-A-0 121 386 - DNA vectors and their Use in Recombinant DNA Technology.

WO-A-8 902 465 - Polypeptide Production.

WO-A-8 909 825 - Method for Producing Recombinant DNA Proteins.

EP-A-0 338 841 - Recombinant DNA Methods, Vectors and Host Cells.

WO-A-9 010 700 - Modified Antibodies.

WO-A-9 008 191 - Recombinant DNA Method.

WO-A-9 106 659 - Recombinant DNA Method and Vectors for use therein.

WO-A-9 119 739 - Multivalent antigen-Binding Proteins.

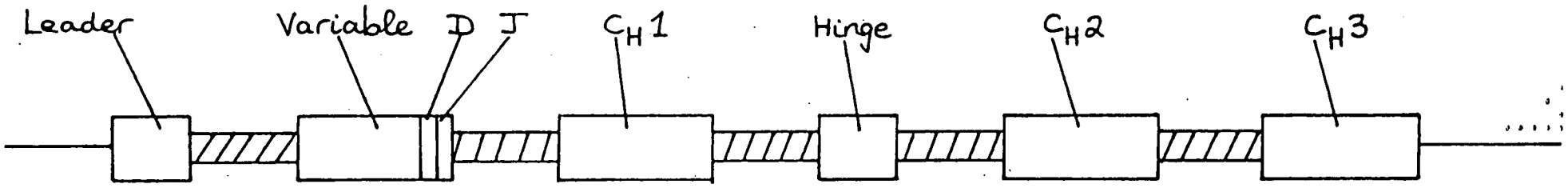
IN THE MATTER of
European Patent No. 0 125 023
of Genentech, Inc.
-and-
IN THE MATTER of an
Opposition thereto filed by
Celltech Limited

I hereby certify that this is Exhibit GTY3 to the Statutory Declaration of Geoffrey Thomas YARRANTON executed this day of 1992.

IN THE MATTER of
European Patent No. 0 125 023
of Genentech, Inc.
-and-
IN THE MATTER of an
Opposition thereto filed by
Celltech Limited

I hereby certify that this is Exhibit GTY4 to the Statutory Declaration of Geoffrey Thomas YARRANTON executed this day of 1992.

Schematic Representation of Rearranged Immunoglobulin Heavy Chain Gene.



- Exons

- Introns

——— - Untranslated Regions

IN THE MATTER of
European Patent No. 0 125 023
of Genentech, Inc.
-and-
IN THE MATTER of an
Opposition thereto filed by
Celltech Limited

I hereby certify that this is Exhibit GTY5 to the Statutory Declaration of Geoffrey Thomas YARRANTON executed this day of 1992.

CONSTRUCTION OF CHIMERIC GENE CONSTRUCTS

Example 1: Using Site-Directed Mutagenesis and DNA Fragment Assembly

Using site directed mutagenesis techniques (Messing et al., Nuc. Acids. Res., 9, 309 (1981)) on a genomic clone of the human γ4 constant region it was possible in 1982 to introduce restriction sites into a genomic human γ4 clone allowing the assembly of an intronless copy of the constant region. The sequence of a human genomic clone of the γ4 constant region was published in 1981 (Ellison et al., DNA, 1, 11-18, 1981). Using an equivalent genomic clone the following manipulations could have been carried out:-

(1) At the 5' end of the CH1 domain, an XmnI cleavage site would have been introduced. A mutagenic oligonucleotide of the sequence: 5' CCTCTCGAACAGCTTCC would have been used for this. Digestion with XmnI generates a blunt end cleavage site before the first codon (GCT) of the CH1 domain.

(2) At the 3' end of the CH1 domain, an XbaI cleavage site would have been introduced. A mutagenic oligonuleotide of the sequence: 5' AAGAGAGTCTAGAAGAGG would have been used for this. Cleavage at the sequence TCTAGA by XbaI would generate a 5' overhang sequence of 5' CTAG. To generate a blunt end after the final codon of the CH1 domain the cleaved DNA would need to be filled in using Klenow DNA polymerase using only dGTP, followed by removal of single stranded ends using SI nuclease.

(3) The hinge region and 5' end of the CH_2 domain would have been assembled from oligonucleotides. A naturally occurring AvaII restriction enzyme site would have been used to join the remaining part of the CH_2 domain to the assembled hinge region. The oligonucleotides used in this synthesis would have been:

```
5'    GAG TCC AAA TAT GGT CCC CCA TGC CCA
      CTC AGG TTT ATA CCA GGG GGT ACG GGT AGT ACG   5'

5'    TCA TGC CCA GCA CCT GAG TTC CTG GGG G
            GGT CGT GGA CTC AAG GAC CCC CCT G   5'
```

(4) Close to the 3' end of the CH_2 domain, a ClaI restriction site would have been introduced using a mutagenic oligonucleotide of the sequence: 5' CTCCCGTCATCGATCGAG. This would have allowed the isolation of most of the CH_2 domain on an AvaII to ClaI DNA fragment.

(5) Close to the 5' end of the CH_3 domain sequence, a BamHI restriction site would have been introduced using a mutagenic oligonucleotide of the sequence: CAGCCCCGGGATCCACAGGTG

(6) The 3' end of the CH_2 domain would have been joined from the created ClaI site to the 5' end of the CH_3 domain at the created BamHI site using the oligonucleotides:

```
5'    CGA TCG AGA AAA CCA TCT CCA AAG CC
        T AGC TCT TTT GGT AGA GGT TTC GG TTT     5'

5'    AAA GGG CAG CCC CGG
          CCC GTC GGG GCC CTA G    5'
```

(7) A HindIII restriction site would have been introduced into the gene sequence following the translational stop codon (TGA) using the mutagenic oligonucleotide 5' GGCGGCAAGCTTCCGCTCC. This would have allowed the isolation of a DNA fragment (BamHI to HindIII) comprising most of the CH3 domain sequence.

Using standard recombinant DNA techniques, the human γ4 constant region gene would have been assembled as an XmnI-HindIII DNA fragment using the scheme devised above. Cleavage at the XmnI site would generate a blunt end site for ligation of any cDNA copy of an immunoglobulin variable domain gene sequence. For example the variable domain gene for the B1-8 antibody belonging to the NP family could have been used. Antibodies of this family bind to the hapten (4-hydroxy-3-nitrophenyl) acetyl. The sequence of this variable domain gene was published in 1982 (Bothwell et al., Cell, 24, 623-637, 1982). The 360 base pair sequence could have been chemically synthesised. Using oligonucleotides of the sequence 5' AGATCTATGGGATGG at the 5' end of the coding sequence would have provided a BglII restriction site prior to the sequence for the intiation codon (ATG). Ligation to the human γ4 constant region at the blunt ended XmnI restriction site would have been achieved by blunt end ligation at the 3' end of the assembled variable domain gene. The assembled mouse-human chimeric gene would then have been isolated as a BglII to HindIII restriction fragment.

Chimeric light chain gene constructs would have been made using similar methodology, since the human kappa light chain constant region gene sequence was published in 1980 (Hieter et al., Cell, 22, 197-207).

Hence by 1982, recombinant DNA technology techniques and DNA sequence information were available such that chimeric gene constructs could have been prepared for expression studies in microbial or mammalian host cells.

Example 2: Gene Assembly of a Human Constant Region Gene

Recombinant DNA technology had been used in the early nineteen eighties to assemble genes completely from chemically synthesised oligonucleotides (e.g. α-interferon, Edge et al., Nature (London), 292, 756-762, 1981). Using these techniques the human constant region could have been assembled from the published genomic DNA sequences. For example, the human γ4 constant region gene could have been assembled as an Xmnl-HindIII fragment, using the oligonucleotides listed below. The gene would have been assembled in a stepwise manner using the restriction enzyme sites indicated. The final DNA vector would have comprised a unique XmnI cleavage site for the insertion of a cloned or chemically synthesised variable region gene sequence.

The light chain human kappa gene sequence could have been similarly assembled.

GENE SYNTHESIS FOR HUMAN γ4 CONSTANT REGION

XmnI

1

5' GAANN GCT,TCC,ACC,AAG,GGC,CCA,TCC,GTC,TTC
CGA AGG,TGG,TTC,CCG,GGT,AGG,CAG AAGGGG GAC

```
                                    3
5'      CC,C,CTG,GCG,CCC TGC TCC AGG AGC ACC TCC GAG AGC
                 CGC GGG ACG AGG TCC TCG TGG AGG CTC TCG,TGT,CGG
                                    4

                                          5                                    HpaI
5'      ACA GCC GCC CTG GGC TGC CTG GTC AAG GAC TAC TTC CCC GAACCGGTT
                 CGG GAC CCG ACG GAC CAG TTC CTG ATG AAG GGG CTTGGCCAA
                                          6


                                             7
5'      ACC GTG TCG TGG AAC TCA GGC GCC CTG ACC AGC GGC
        TGG CAC AGC ACC TTG AGT CCG CGG GAC TGG TCG CCGCACGTG
                                             8

                                          9
5'      GTG CAC ACC TTC CCG GCT GTC CTA CAG T CCT CAG GAC TCT
                 TGG AAG GGC CGA CAG GAT GTC A GGA GTC CTG AGATGAGGG
                                          10

                                             11
5'      ACT CCC TCA GCA GCG TGG TGA CCG TGC CCT CCA GCA GCT
                 AGT CGT CGC ACC ACT GGC ACG GGA GGT CGT CGAACCCGT
                                             12

                                       13           SalI
5'      T GGGCACG,AAG,ACC,TAC ACC TGC AAC,G
                 G CTT CTG GAT GTG GAC GTT GCAGCT
                                       14

         SalI                             15
5'      TC GAC CAC AAG CCC AGC AAC ACC AAG GTG GAC AAG
             G GTG TTC GGG TCG TTG TGG TTC CAC CTG TTC TCT CAA
                                          16

                                       17
5'      AGA GTT GAG,TCC,AAA TAT GGT CCC CCA TGC CCA TCA
                 CTC AGG TTT ATA CCA GGG GGT ACG GGT AGT ACG GGT
                                       18

                                          19
5'      TGC CCA GCA CCT GAG TTC CTG GGG GGA CCA TCA
                 CGT GGA CTC AAG GAC CCC CCT GGT AGT CAG AAG GAC
                                          20

                                             21
5'      GTC TTC CTG TTC CCC CCA AAA CCC AAG GAC ACT CTC ATG
                 AAG GGG GGT TTT GGG TTC CAG TGA GAG TAC TAG AGG
                                             22
```

```
                                    23
5'      ATC TCC CGG ACC CCT GAG GTC ACG TGC GTG
                GCC TGG GGA CAC CAG TGC ACG CAC CAC CAC
                                    24

                              25                    BamHI
5'      GTG GTG GAC GTG AGC CAG GAG
                CTG CAC TCG GTC CTC CTA G
                               26

BamHI                                 27
5'      GAT CCC GAG GTC CAG TTC AAC TGG TAC GTG GAT
             GGC TCC AGG TCA AGT TGA CCA TGC ACC TAC AGC AC
                                    28


                                                29
5'      GGC GTG GAG GTG CAT AAT GCC AAG ACA AAG CCG CGG
                CTC CAC GTC TTA CGG TTC TGT TTC GGC GCC CTC CTC
                                      30

                                   31
5'      GAG GAG CAG TTC AAC AGC ACG TAC CGT GTG GTC
                GTC AAG TTG TCG TGC ATG GCA CAC CAG TCG CAG
                                    32

                                          33
5'      AGC GTC CTC ACC GTC CTG CAC CAG GAC TGG CTG AAC GGC AAG
                GAG TGG CAG GAC GTG GTC CTG  ACC GAC TTG CCG TTC CTC ATG
                                    34

                                    35                          ClaI
5'      GAG TAC AAG TGC AAG GTC TCC AAC AAA GGC CTC CCG TCA T
                TTC ACG TTC CAG AGG TTG TTT CCG GAG GGC AGT AGC
                                    36

    ClaI                                37
5'    CG ATC GAG AAA ACC ATC TCC AAA GCC AAA GGG CAG CCC
         TAG CAC TTT TGG TAG AGG TTT CGG TTT CCC GTC GGG GCT CTC
                                    38

                                             39
5'      CGA GAG CCA CAG GTG TAC ACC CTG CCC CCA TCC CAG GAG GAG
                GGT GAC CAC ATG TGG GAC GGG GGT AGG GTC CTC CTC TAC TGG
                                       40

                                    41
5'      ATG ACC AAG AAC CAG GTC AGC CTG ACC TGC CTG GTC AAA
                TTC TTG GTC CAG TCG GAC TGG ACG GAC CAG TTT CCG AAG
                                    42

                                          43
5'      GGC TTC TAC CCC AGC GAC ATC GCC GTG GAG TGG GAG
                ATG GGG TCG CTG TAG CGG CAC CTC ACC CTC TCG TTA
                                    44
```

```
5'      AGC AAT GGG CAG CCG GAG AAC AAC TAC AAG ACC ACG
               CCC GTC GGC CTC TTG TTG ATG TTC TGG TGC GGA GGG
                                    46

                                    47
5'      CCT CCC GTG CTG GAC TCC GAC GGC TCC TTC TTC CTC
               CAC GAC CTG AGG CTG CCG AGG AAG AAG GAG ATG TCG
                                         48

                          49          SalI
5'      TAC AGC AGG CTA ACC G
               TCC GAT TGG CAGCT
                          50


SalI                                  51
5'      TC GAC AAG AGC AGG TGG CAG GAG GGG AAT GTC TTC
             G TTC TCG TCC ACC GTC CTC CCC TTA CAG AAG AGT ACG
                                    52

                                    53
5'      TCA TGC TCC GTG ATG CAT GAG GCT CTG CAC AAC CAC
               AGG CAC TAC GTA CTC CGA GAC GTG TTG GTG ATG TGT
                                    54

                               55                                   HindIII
5'      TAC ACA CAG AAG AGC CTC TCC CTG TCT CTG GGT AAA TGA
               GTC TTC TCG GAG AGG GAC AGA GAC CCA TTT ACT TCGA
                               56
```

As described above, the chemically synthesised human γ4 constant region gene would have been ligated to a blunt-ended variable region gene to provide a chimeric construct. The BglII to blunt ended DNA fragment described for B1-8 would have provided such a sequence.